

Min-Hung (Steve) Chen 陳敏弘

Email: vitec6@gmail.com | Website: minhungchen.netlify.app/ | LinkedIn: [chensteven](https://www.linkedin.com/in/chensteven) | GitHub: [cmhungsteve](https://github.com/cmhungsteve)

RESEARCH INTERESTS

Multimodal AI, Efficient AI, Video Understanding, Vision Transformer, Transfer Learning, Computer Vision, Deep Learning

EDUCATION

Ph.D., Georgia Institute of Technology (GPA: 4.00/4.00)

August 2014 – August 2020

School of Electrical and Computer Engineering, College of Engineering

Atlanta, U.S.A.

Thesis topic: *Bridging Distributional Discrepancy with Temporal Dynamics for Video Understanding*

RESEARCH WORK EXPERIENCE

Senior Research Scientist, NVIDIA Research, Supervisor: Prof. Yu-Chiang Frank Wang

November 2022 – Present

- Vision+X Multimodal AI: Vision-Language, 4D (video+depth) Understanding, VLA, Efficient AI, [Transformer](#), etc.
- [ICML'24 Oral] DoRA: a default replacement of LoRA ([paper](#), [code](#), [website](#), [NV Tech Blog](#), [Hugging Face](#))
- EoRA: a fine-tuning-free approach for rapidly recovering LLM compression errors ([paper](#), [code](#), [NV Tech Blog](#))

Research Engineer II, Microsoft, Supervisor: Prof. Shang-Hong Lai

January 2022 – October 2022

- Conducted research and deployment for generalizable and explainable facial liveness approaches for Azure Cloud AI.

Senior AI Engineer, MediaTek Inc., Supervisor: Dr. Yi-Min Tsai

October 2020 – December 2021

- Edge-AI research (e.g., on-device Transformer) and academic-industry collaboration (e.g., co-host CVPR' 21 workshop).

Ph.D. Research, OLIVES, Advisor: Prof. Ghassan AlRegib

August 2014 – August 2020

- [CVPR'20, ICCV'19 (Oral)] Video domain adaptation, [SPIC'19] Action recognition, [TITS'19] Large traffic sign dataset
- Research internships: Baidu USA (Summer'19), PlayStation (Summer'18), V7 Labs (Fall'17)

Research Assistant, Academia Sinica, Supervisor: Dr. Yen-Yu Lin

July 2013 – July 2014

- [CVPR'14] Developed transfer learning techniques to leverage depth and skeleton knowledge for action recognition.

SELECTED PUBLICATIONS (*equal contribution / †project lead)

([full paper list](#))

- Ying Cheng, Yu-Ho Lin, **Min-Hung Chen**, Fu-En Yang, and Shang-Hong Lai. “[VADER: Towards Causal Video Anomaly Understanding with Relation-Aware Large Language Models](#)”, *WACV*, 2026.
- Chi-Pin Huang, Yueh-Hua Wu, **Min-Hung Chen**, Yu-Chiang Frank Wang, and Fu-En Yang. “[ThinkAct: Vision-Language-Action Reasoning via Reinforced Visual Latent Planning](#)”, *NeurIPS*, 2025.
- Jin-Ting He, Fu-Jen Tsai, Yan-Tsung Peng, **Min-Hung Chen**, Chia-Wen Lin, and Yen-Yu Lin. “[BlurDM: A Blur Diffusion Model for Image Deblurring](#)”, *NeurIPS*, 2025.
- Gueter Josmy Faure, **Min-Hung Chen**, Jia-Fong Yeh, Ying Cheng, Hung-Ting Su, Yung-Hao Tang, Shang-Hong Lai, and Winston H. Hsu. “[MovieCORE: COgnitive REasoning in Movies](#)”, *EMNLP*, 2025. [Oral]
- Gueter Josmy Faure, Jia-Fong Yeh, **Min-Hung Chen**, Hung-Ting Su, Shang-Hong Lai, and Winston H. Hsu. “[HERMES: temporal-coHERent long-forM understanding with Episodes and Semantics](#)”, *ICCV*, 2025.
- Chin-Yang Lin, Cheng Sun, Fu-En Yang, **Min-Hung Chen**, Yen-Yu Lin, and Yu-Lun Liu. “[LongSplat: Robust Unposed 3D Gaussian Splatting for Casual Long Videos](#)”, *ICCV*, 2025.
- Hsu-kuang Chiu, Ryo Hachiuma, Chien-Yi Wang, Stephen F. Smith, Yu-Chiang Frank Wang, and **Min-Hung Chen**†. “[V2V-LLM: Vehicle-to-Vehicle Cooperative Autonomous Driving with Multi-Modal Large Language Models](#)”, *CVPRW*, 2025.
- Miran Heo, **Min-Hung Chen**, De-An Huang, Sifei Liu, Subhashree Radhakrishnan, Seon Joo Kim, Yu-Chiang Frank Wang, and Ryo Hachiuma. “[Omni-RGPT: Unifying Image and Video Region-level Understanding via Token Marks](#)”, *CVPR*, 2025.

- Chung-Ho Wu, Yang-Jung Chen, Ying-Huan Chen, Jie-Ying Lee, Bo-Hsu Ke, Chun-Wei Tuan Mu, Yi-Chuan Huang, Chin-Yang Lin, **Min-Hung Chen**, Yen-Yu Lin, and Yu-Lun Liu. “[AuraFusion360: Augmented Unseen Region Alignment for Reference-based 360° Unbounded Scene Inpainting](#)”, *CVPR*, 2025.
- Xin Dong, Yonggan Fu, Shizhe Diao, Wonmin Byeon, Zijia Chen, Ameya Sunil Mahabaleshwarkar, Shih-Yang Liu, Matthijs Van Keirsbilck, **Min-Hung Chen**, Yoshi Suhara, Yingyan (Celine) Lin, Jan Kautz, and Pavlo Molchanov. “[Hymba: A Hybrid-head Architecture for Small Language Models](#)”, *ICLR*, 2025. [Spotlight]
- Yusuke Hirota, **Min-Hung Chen**, Chien-Yi Wang, Yuta Nakashima, Yu-Chiang Frank Wang, and Ryo Hachiuma. “[SANER: Annotation-free Societal Attribute Neutralizer for Debiasing CLIP](#)”, *ICLR*, 2025.
- Ci-Siang Lin, Chien-Yi Wang, Yu-Chiang Frank Wang, and **Min-Hung Chen**[†]. “[Semantic Prompt Learning for Weakly-Supervised Semantic Segmentation](#)”, *WACV*, 2025. [accepted in Round 1]
- Chun-Mao Lai, Hsiang-Chun Wang, Ping-Chun Hsieh, Yu-Chiang Frank Wang, **Min-Hung Chen**, and Shao-Hua Sun. “[Diffusion-Reward Adversarial Imitation Learning](#)”, *NeurIPS*, 2024.
- Shih-Yang Liu, Chien-Yi Wang, Hongxu Yin, Pavlo Molchanov, Yu-Chiang Frank Wang, Kwang-Ting Cheng, and **Min-Hung Chen**[†]. “[DoRA: Weight-Decomposed Low-Rank Adaptation](#)”, *ICML*, 2024. [Oral (acceptance rate: 1.5%)]
- Ardian Umam, Cheng-Kun Yang, **Min-Hung Chen**, Jen-Hui Chuang, and Yen-Yu Lin. “[PartDistill: 3D Shape Part Segmentation by Vision-Language Model Distillation](#)”, *CVPR*, 2024.
- Ji-Jia Wu, Chia-Hao Chang, Chieh-Yu Chuang, Chun-Pei Chen, Yu-Lun Liu, **Min-Hung Chen**, Hou-Ning Hu, Yung-Yu Chuang, and Yen-Yu Lin. “[Image-Text Co-Decomposition for Text-Supervised Semantic Segmentation](#)”, *CVPR*, 2024.
- Cheng-Kun Yang, **Min-Hung Chen**, Yung-Yu Chuang, and Yen-Yu Lin. “[2D-3D Interlaced Transformer for Point Cloud Segmentation with Scene-Level Supervision](#)”, *ICCV*, 2023.
- Su-Kai Chen, Hung-Lin Yen, Yu-Lun Liu, **Min-Hung Chen**, Hou-Ning Hu, Wen-Hsiao Peng, and Yen-Yu Lin. “[Learning Continuous Exposure Value Representations for Single-Image HDR Reconstruction](#)”, *ICCV*, 2023.
- **Min-Hung Chen**, Baopu Li, Yingze Bao, Ghassan AlRegib, and Zsolt Kira. “[Action Segmentation with Joint Self-Supervised Temporal Domain Adaptation](#)”, *CVPR*, 2020.
- **Min-Hung Chen**, Zsolt Kira, Ghassan AlRegib, Jaekwon Yoo, Ruxin Chen, and Jian Zheng. “[Temporal Attentive Alignment for Large-Scale Video Domain Adaptation](#)”, *ICCV*, 2019. [Oral (acceptance rate: 4.6%), travel grant awarded]
- Chih-Yao Ma^{*}, **Min-Hung Chen**^{*}, Zsolt Kira and Ghassan AlRegib. “[TS-LSTM and Temporal-Inception: Exploiting Spatiotemporal Dynamics for Activity Recognition](#)”, *Elsevier SPIC*, 2019. [2023 EURASIP Best Paper Award]
- Dogancan Temel, **Min-Hung Chen**, and Ghassan AlRegib. “[Traffic Sign Detection Under Challenging Conditions: A Deeper Look into Performance Variations and Spectral Characteristics](#)”, *IEEE T-ITS*, 2019.

PROFESSIONAL ACTIVITIES

- Area Chairs: *ICLR*, *ICML*
- Workshop Organizers: *ICCV* ([X-Sense](#)), *CVPR* ([T4V](#)), *CVPR* ([MAI](#)), *AIVR*
- Professional Talks: *CVPR*, *ICCV*, *Academia Sinica*, *NYCU*, *NTHU*, *NTU*
- Conference Reviewers: *CVPR*, *ICLR*, *NeurIPS*, *ICCV*, *ICML*, *ECCV*, *AAAI*, *WACV*, *BMVC*, *ICIP*, *ACCV*, *ICASSP*, etc.
- Journal Reviewers: *IEEE T-PAMI*, *Pattern Recognition*, *IJCV*, *IEEE T-ITS*, *IEEE T-CSVT*, *IEEE Access*

HONORS AND AWARDS

- Outstanding Reviewers: [ICLRW 2025](#), [ICML 2022](#), [ICCV 2021](#), [CVPR 2021](#)
- [EURASIP Best Paper Award](#) for Image Communication Journal Fall 2023
- Ministry of Education Technologies Incubation Scholarship, Taiwan (R. O. C.) Fall 2014 – Spring 2017

SKILLS

- **Technical:** PyTorch / Python / Tensorflow / LuaJIT / Torch / Caffe / OpenCV / C++ / MATLAB / Linux / Unreal Engine
- **Languages:** Mandarin Chinese / English / Japanese (JLPT Level N2 Passed)

TEACHING EXPERIENCE

Graduate Teaching Assistant, Georgia Institute of Technology August 2014 – May 2019

- Deep Learning (Spring 2019); Computer Vision (Fall 2018)
- Signals and Systems (Spring 2015); Fundamentals of Digital Signal Processing (Fall 2014)

Graduate Teaching Assistant, National Taiwan University September 2011 – June 2012

- Statistical Image Processing (Spring 2012); Computer Programming (Fall 2011)

RELATED COURSE PROJECTS

Deep Learning for Video Classification; Character Recognition in Natural Images; Gender Classification by Face Analysis; Light Field Camera Refocusing; Multi-label Classification with Missing Data.

EXTRACURRICULAR ACTIVITIES

In Georgia Tech:

- President, Taiwanese Student Association in Georgia Tech Fall 2015 – Spring 2016

International Summer Program:

“Asia in Today’s World” Program, Kyushu University Summer 2011

- Final project topic: Religion Influence on Japan’s Robot Development.

American Language & Culture Program, Stanford University Summer 2007

- Final project topic: Individualism in America.

Others:

Department Basketball Team (6 years); Delicious Club (2 years); Service Volunteer; NTUEE Annual Performance; International Business Camp; NTUEE Summer Camp; Seminar on International Trade and Economy